AD-776 610

# THE VALUE OF SEQUENTIAL INFORMATION

Allen Clinton Miller, III

Stanford University

Prepared for:

Office of Naval Research
Advanced Research Projects Agency
National Science Foundation

31 January 1974

Research Report No. EES-DA-73-2
January 1974

# THE VALUE OF SEQUENTIAL INFORMATION

ALLEN C. MILLER III

DECISION ANALYSIS PROGRAM
Professor Ronald A. Howard
Principal Investigator



DEPARTMENT OF ENGINEERING-ECONOMIC SYSTEMS
Stanford University
Stanford, California 94305



SPONSORSHIPS

Advanced Research Projects Agency, Human Resources Research Office ARPA Order No. 2449, monitored by Engineering Psychology Programs, Office of Naval Research, under Contract No. N00014-67-A-0112-0077 (NR 197-024) covering the period May 1, 1973 to December 30, 1973.

National Science Foundation under NSF Grant GK-36491.

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

| REPORT DOCUMENTATION PAGE | | READ INSTRUCTIONS BEFORE COMPLETING FORM |
|---|---|---|
| 1. REPORT NUMBER<br>EES-DA-73-2 | 2. GOVT ACCESSION NO. | 3. RECIPIENT'S CATALOG NUMBER |
| 4. TITLE (and Subtitle)<br>"The Value of Sequential Information" | | 5. TYPE OF REPORT & PERIOD COVERED<br>Technical<br>5/1/73 to 12/30/73 |
| | | 6. PERFORMING ORG. REPORT NUMBER |
| 7. AUTHOR(s)<br>Allen Clinton Miller, III | | 8. CONTRACT OR GRANT NUMBER(s)<br>N00014-67-A-0112-0077 |
| 9. PERFORMING ORGANIZATION NAME AND ADDRESS<br>The Board of Trustees of the Leland Stanford Junior University, c/o Office of Research Administrator, Encina Hall, Stanford, California 94305 | | 10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS<br>000000<br>ARPA Order #2449 |
| 11. CONTROLLING OFFICE NAME AND ADDRESS<br>Advanced Research Projects Agency<br>Human Resources Research Office<br>1400 Wilson Blvd., Arlington, Virginia 22209 | | 12. REPORT DATE<br>January 31, 1974 |
| | | 13. NUMBER OF PAGES<br>24 pages |
| 14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)<br>Engineering Psychology Programs, Code 455<br>Office of Naval Research<br>800 N. Quincy Street, Arlington, VA 22217 | | 15. SECURITY CLASS. (of this report)<br>Unclassified |
| | | 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE |
| 16. DISTRIBUTION STATEMENT (of this Report)<br>Approved for public release; distribution unlimited. | | |
| 17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report) | | |
| 18. SUPPLEMENTARY NOTES<br>A paper submitted to MANAGEMENT SCIENCE/THEORY. | | |
| 19. KEY WORDS (Continue on reverse side if necessary and identify by block number)<br>DECISION ANALYSIS · DECISION THEORY · SEQUENTIAL INFORMATION<br>DECISION-MAKING · INFORMATION · VALUE OF INFORMATION | | |

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

In decision analysis we normally consider the value of information to be a constant against which the cost of information is compared. However, when it is possible to buy information sequentially, the value of information is not a constant. Rather, it is a function of the prices of the various pieces of information, or "observables." When we are faced with a decision and learn one observable, this information not only helps us make the

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102-014-6601

UNCLASSIFIED
SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

ii

Block 20 (continued)

original decision, but also helps us decide if we should pay for more observables. For this reason, the first observable has a value above and beyond that which we would assign if there were no possibility of obtaining additional information. To decide whether or not to buy one observable we must know the prices of all the observables.

# THE VALUE OF SEQUENTIAL INFORMATION*

Allen C. Miller, III

## Abstract

In decision analysis we normally consider the value of information to be a constant against which the cost of information is compared. However, when it is possible to buy information sequentially, the value of information is not a constant. Rather, it is a function of the prices of the various pieces of information, or "observables." When we are faced with a decision and learn one observable, this information not only helps us make the original decision, but also helps us decide if we should pay for more observables. For this reason, the first observable has a value above and beyond that which we would assign if there were no possibility of obtaining additional information. To decide whether or not to buy one observable we must know the prices of all the observables.

## Introduction

Using decision analysis [1,4,6,7]** it is possible to calculate the value of one or more pieces of information--called "observables"--when a decision must be made in the face of uncertainty. This information has value because it can affect the decision and lead to a greater expected

*This research was partially supported by the Advanced Research Projects Agency of the Department of Defense, as monitored by the Office of Naval Research under Contract No. N00014-67-A-0112-0077, and by the National Science Foundation under NSF Grant GK-36491.

**Numbers in square brackets refer to the References.

profit. However the possibility of buying information sequentially presents the decision maker with a set of secondary decisions: which observables should he buy and in which order should he buy them? It is possible that knowing one observable affects not only the primary decision, but also the decision to buy additional information. In that case the value of knowing the first observable is greater than it would be if it affected only the primary decision. The prices of the observables affect the decision maker's willingness to buy additional information. For this reason the amount that the value of learning each obserbable is increased by the possibility of buying additional information depends on the prices of all the observables.

## Notation

The notation used in this paper is an extension of Howard's inferential notation [2,3], which shows the decision maker's state of information explicitly. Although inferential notation is very useful, the following operator notation is somewhat clearer and more compact when dealing with the value of information.

$\mathcal{S}$ = a state of information on which probability assignments are made,

$\mathcal{E}$ = the decision maker's prior state of information,

$\{x|\mathcal{S}\}$ = the density function (for continuous random variables) or mass function (for discrete random variables) of the random variable $x$, given the state of information $\mathcal{S}$,

$\langle x|\mathcal{S}\rangle = \int x\{x|\mathcal{S}\}\,dx$ = the expected value of $x$ given the state of information $\mathcal{S}$,

$\bar{x} = \langle x|\mathcal{S}\rangle$ = the expected value of $x$ given the decision maker's prior state of information ,

$\pi(x_1, \ldots, x_n, c)$ = a profit function that depends on the state and control variables,

$c$ = a control or decision variable (possibly vector-valued) upon which the profit depends ,

$x_i$ = a state variable; a random variable upon which the profit depends. (If we are dealing with perfect information, $x_i$ is also an observable.)

$y_j$ = an observable; a random variable whose actual value we have an opportunity to learn ,

$\underset{x_i}{E}\, \pi(x_1, \ldots, x_n, c) = \int \pi(x_1, \ldots, x_n, c) \left\{x_i|\mathcal{S}\right\} dx$ = the expected value of the profit with respect to $x_i$ ,

$\max_c \pi(x_1, \ldots, x_n, c)$ = the maximum of the profit with respect to $c$ ,

$K_{x_i}$ = cost of learning $x_i$ ,

$V^N_{x_i}$ = the value of learning $x_i$ given that no additional information is purchased; the value of individual information about $x_i$ (the superscript $N$ means "no additional information") ,

$V^N_{x_i x_j \ldots x_r}$ = the value of learning $x_i$ , $x_j$ , ..., and $x_r$ simultaneously given that no additional information is purchased; the value of simultaneous information about $x_i$ , $x_j$ , ..., and $x_r$ ,

$V^R_{x_i} = V^N_{x_1 \ldots x_n} - K_{x_1} - \ldots - K_{x_{i-1}} - K_{x_{i+1}} - \ldots - K_{x_n}$, the residual value of learning $x_i$ when all of the available information is purchased simultaneously; the residual value of $x_i$, (the superscript R stands for "residual value"),

$V_{x_i}$ = the value of learning $x_i$ when additional information can be purchased; the value of sequential information about $x_i$.

Strings of expectation and maximization operators are interpreted as follows:

$$\max_c \mathop{E}_x \pi(x,c) = \max_c \left( \int_{-\infty}^{\infty} \pi(x,c) \; \{x|\mathcal{E}\} \; dx \right)$$

$$\mathop{D}_y \max_c \mathop{E}_x \pi(x,c) = \max_c \left( \int_{-\infty}^{\infty} \pi(x,c) \; \{x|y,\mathcal{E}\} \; dx \right)$$

$$\mathop{E}_y \max_c \mathop{E}_x \pi(x,c) = \int_{-\infty}^{\infty} \max_c \left( \int_{-\infty}^{\infty} \pi(x,c) \; \{x|y,\mathcal{E}\} \; dx \right) \{y|\mathcal{E}\} \; dy$$

$\mathop{D}_y$ is a modifier that conditions the expectations in its argument on $y$. $\mathop{E}_x$ is interpreted as the expected value with respect to $x$ given all the control or random variables that appear as subscripts of operators or modifiers that include $\mathop{E}_x$ in their argument.

In this paper we assume that none of the probability density functions depends on the control variable.

<u>Types of Sequential Information Problems</u>

A decision problem involving sequential information can be placed in one of several categories depending on whether or not the problem has each of the following properties.

1. Additive prices. The prices of the observables are additive

when the cost of learning any group of observables simultaneously equals the sum of the costs of learning each observable separately.

2. Certain prices. If the decision maker knows the price of each observable and group of observables, the problem can be analyzed in terms of certain prices.

3. Perfect information. When a piece of information reveals the exact value of one or more of the arguments of the profit function, it is called perfect information.

For simplicity, we do not consider utility and risk preference in this paper. However the extension of the following results to include these concepts is straightforward.

The following example introduces ideas of sequential information with the simplest type of problem: the case of additive, certain prices for perfect information. After the example the results are generalized to apply to all types of sequential information problems.

## An Example: A Bidding Problem Revisited

The example addresses the problem of submitting a single, sealed bid for a contract when faced with uncertainty about the lowest competing bid and the production cost required to fulfill the contract. This problem is discussed in some detail by Howard [2,3].

Suppose our company's objective in bidding is to maximize its expected profit. Let $p$ be our production cost for performing the contract, let $\ell$ be our competitors' lowest bid, and let $b$ be our company's bid. If our bid exceeds the lowest competing bid, our company

will not win the contract and our net profit will be zero. However if we submit the lowest bid, our company will win the contract and make a profit equal to the difference between our bid and the production cost. Thus our company's profit, $\pi$, is defined by

$$\pi(p,\ell,b) = \begin{cases} b - p : b < \ell \\ 0 \quad : b \geq \ell \end{cases}$$

We do not know the exact values of $p$ and $\ell$, but we can assign prior probability density functions -- $\{p|\mathcal{E}\}$ and $\{\ell|\mathcal{E}\}$ -- based on our company's experience with similar bidding situations. To simplify the computations required for this example we shall assume that $\{p|\mathcal{E}\}$ and $\{\ell|\mathcal{E}\}$ have been assessed as independent, uniform density functions, with $p$ ranging from zero to one and $\ell$ ranging from zero to two. These distributions are shown in Fig. 1.

Consider the possibility of buying perfect information about $p$ and $\ell$ at costs of $K_p$ and $K_\ell$ respectively. Howard [2] has shown that the values of individual and simultaneous information about $p$ and $p$ and $\ell$ are :

$$V_p^N = (\underset{p}{E}\ \underset{b}{\max}\ \underset{\ell}{E} - \underset{b}{\max}\ \underset{p}{E}\ \underset{\ell}{E})\ \pi(p,\ell,b) = 1/96$$

$$V_\ell^N = (\underset{\ell}{E}\ \underset{b}{\max}\ \underset{p}{E} - \underset{b}{\max}\ \underset{p}{E}\ \underset{\ell}{E})\ \pi(p,\ell,b) = 27/96$$

$$V_{p\ell}^N = (\underset{p}{E}\ \underset{\ell}{E}\ \underset{b}{\max} - \underset{b}{\max}\ \underset{p}{E}\ \underset{\ell}{E})\ \pi(p,\ell,b) = 29/96$$

Howard points out that $V_{p\ell}^N$ does not equal the sum of $V_p^N$ and $V_\ell^N$. In other words, the value of learning both $p$ and $\ell$ is greater than the sum of the values of learning $p$ and $\ell$ individually.*

---

*It can be shown that, in general, the value of learning several observables simultaneously can be greater than or less than the sum of the values of learning each observable individually [5].

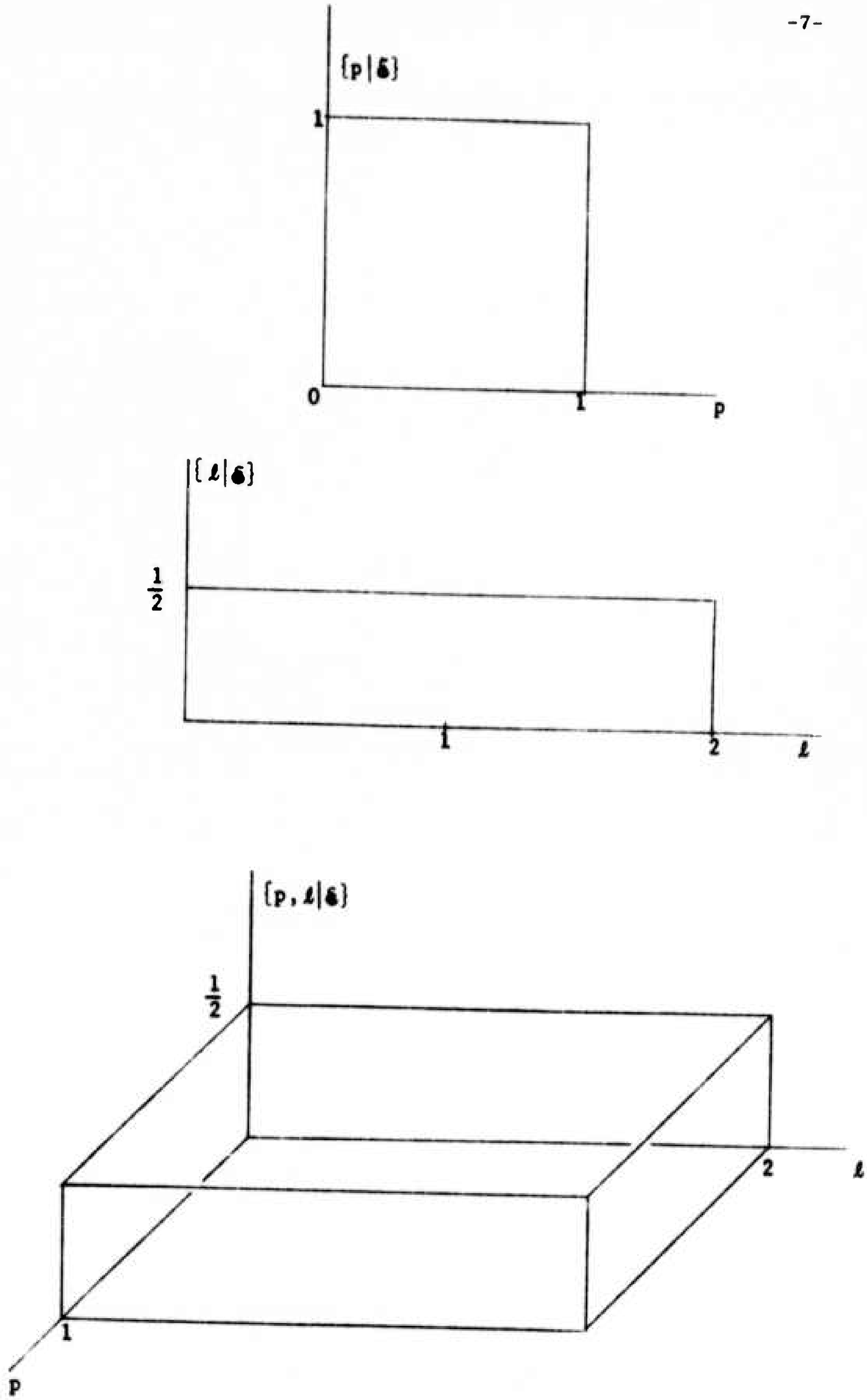


Figure 1. Probability density functions for production cost and lowest competing bid

By comparing the expected profits associated with learning $p$ and $\ell$ individually and simultaneously, we find the decision regions shown in Fig. 2. As Fig. 2 shows, it is necessary to know the prices of both pieces of information to decide which observable to buy, even when only individual and simultaneous purchases are possible. When we consider sequential information it will be possible to draw a diagram similar to Fig. 2, but the decision regions become more complicated.

To determine the value of sequential information about $p$, assume for a moment that we already know $p$ and are trying to decide whether or not to pay to learn $\ell$. If we decide not to learn $\ell$, then our maximum expected profit is:

$$\max_b \mathop{E}_\ell \pi(p,\ell,b) = (1 - p/2)^2/2$$

On the other hand, if we decide to learn $\ell$, our maximum expected profit is

$$\mathop{E}_\ell \max_b \pi(p,\ell,b) = (1 - p/2)^2$$

The increase in expected profit caused by learning $\ell$ after we know $p$ is

$$(\mathop{E}_\ell \max_b - \max_b \mathop{E}_\ell) \pi(p,\ell,b) = (1 - p/2)^2/2$$

We are willing to pay for $\ell$ when the increase in expected profit exceeds $K_\ell$. For certain values of $K_\ell$, the decision to buy perfect information about $\ell$ depends on the value of $p$ that we learned earlier, which means that learning $p$ can help us decide whether to buy more information as well as how to bid. The maximum expected profit when we know that will receive perfect information about $p$ with an option to pay $K_\ell$ for $\ell$ is

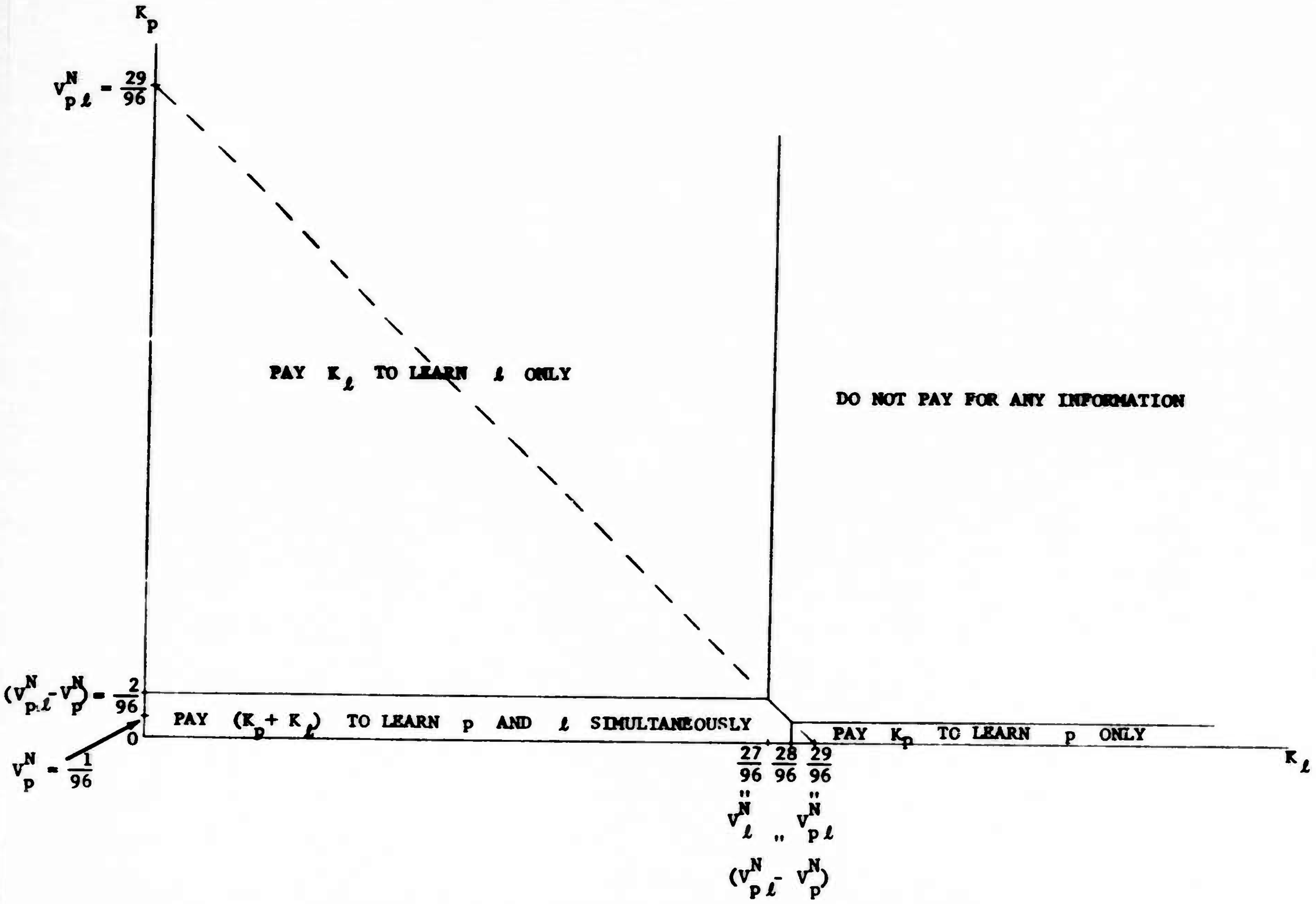


Figure 2. Decision regions for individual and simultaneous information

$$\underset{p}{E}\max\begin{Bmatrix}\underset{\ell}{E}\max_{b}\pi(p,\ell,b)-K_{\ell}\\ \\ \max_{b}\underset{\ell}{E}\pi(p,\ell,b)\end{Bmatrix}$$

$$=\begin{Bmatrix}\underset{p}{E}(1-p/2)^{2}-K_{\ell}:K_{\ell}\leq 1/8\\ \int_{0}^{2(1-\sqrt{2K_{\ell}})}[(1-p/2)^{2}-K_{\ell}]\{p|\delta\}\,dp\\ \qquad+\int_{2(1-\sqrt{2K_{\ell}})}^{1}[(1-p/2)^{2}/2]\{p|\delta\}\,dp:1/8<K_{\ell}<1/2\\ \underset{p}{E}(1-p/2)^{2}/2:1/2\leq K_{\ell}\end{Bmatrix}$$

$$=\begin{Bmatrix}56/96-K_{\ell}:K_{\ell}\leq 1/8\\ 60/96+(4/3)K_{\ell}\sqrt{2K_{\ell}}-2K_{\ell}:1/8<K_{\ell}<1/2\\ 28/96:1/2\leq K_{\ell}\end{Bmatrix}$$

The corresponding increase in expected profit is the value of sequential information about $p$.

$$V_{p}(K_{\ell})=\underset{p}{E}\max\begin{Bmatrix}\underset{\ell}{E}\max_{b}\pi(p,\ell,b)-K_{\ell}\\ \\ \max_{b}\underset{\ell}{E}\pi(p,\ell,b)\end{Bmatrix}-\max_{b}\underset{p}{E}\underset{\ell}{E}\pi(p,\ell,b)$$

$$=\begin{Bmatrix}29/96-K_{\ell}:K_{\ell}\leq 1/8\\ 33/96+(4/3)K_{\ell}\sqrt{2K_{\ell}}-2K_{\ell}:1/8<K_{\ell}<1/2\\ 1/96:1/2\leq K_{\ell}\end{Bmatrix}$$

By carrying out a similar calculation we can find the maximum expected profit when we know that we will receive perfectinformation about $\ell$ with an option to pay $K_p$ for $p$. This expected profit is

$$\mathop{E}_{\ell} \max \left\{ \begin{array}{l} \mathop{E}_{p} \max_{b} \pi(p,\ell,b) - K_p \\ \max E\, \pi(p,\ell,b) \end{array} \right\}$$

$$= \left\{ \begin{array}{l} \int_{\sqrt{2K_p}}^{1-\sqrt{2K_p}} (\ell^2/2 - K_p)\{\ell|\varepsilon\} d\ell + \int_{1-\sqrt{2K_p}}^{2} (\ell - 1/2)\{\ell|\varepsilon\} d\ell \; : \; K_p < 1/8 \\ \int_{1/2}^{2} (\ell - 1/2)\, \{\ell|\varepsilon\} d\ell \; : \; K_p \geq 1/8 \end{array} \right\}$$

$$= \left\{ \begin{array}{l} 56/96 + (2/3)K_p\sqrt{2K_p} - K_p/2 \; : \; K_p < 1/8 \\ 54/96 \; : \; K_p \geq 1/8 \end{array} \right\}$$

The corresponding increase in expected profit is the value of sequential information about $\ell$.

$$V_\ell(K_p) = \mathop{E}_{\ell} \max \left\{ \begin{array}{l} \mathop{E}_{p} \max_{b} \pi(p,\ell,b) - K_p \\ \max_{b} \mathop{E}_{p} \pi(p,\ell,b) \end{array} \right\} - \max E\, E\, \pi(p,\ell,b)$$

$$= \left\{ \begin{array}{l} 29/96 + (2/3)K_p\sqrt{2K_p} - K_p/2 \; : \; K_p < 1/8 \\ 27/96 \; : \; K_p \geq 1/8 \end{array} \right\}$$

By comparing the expected profits associated with all the possible ways we can learn $p$ and $\ell$ --individually, simultaneously and sequentially--we find the decision regions shown in Fig. 3. This figure also shows $V_p(K_\ell)$ and $V_\ell(K_p)$ plotted against the appropriate axes, since

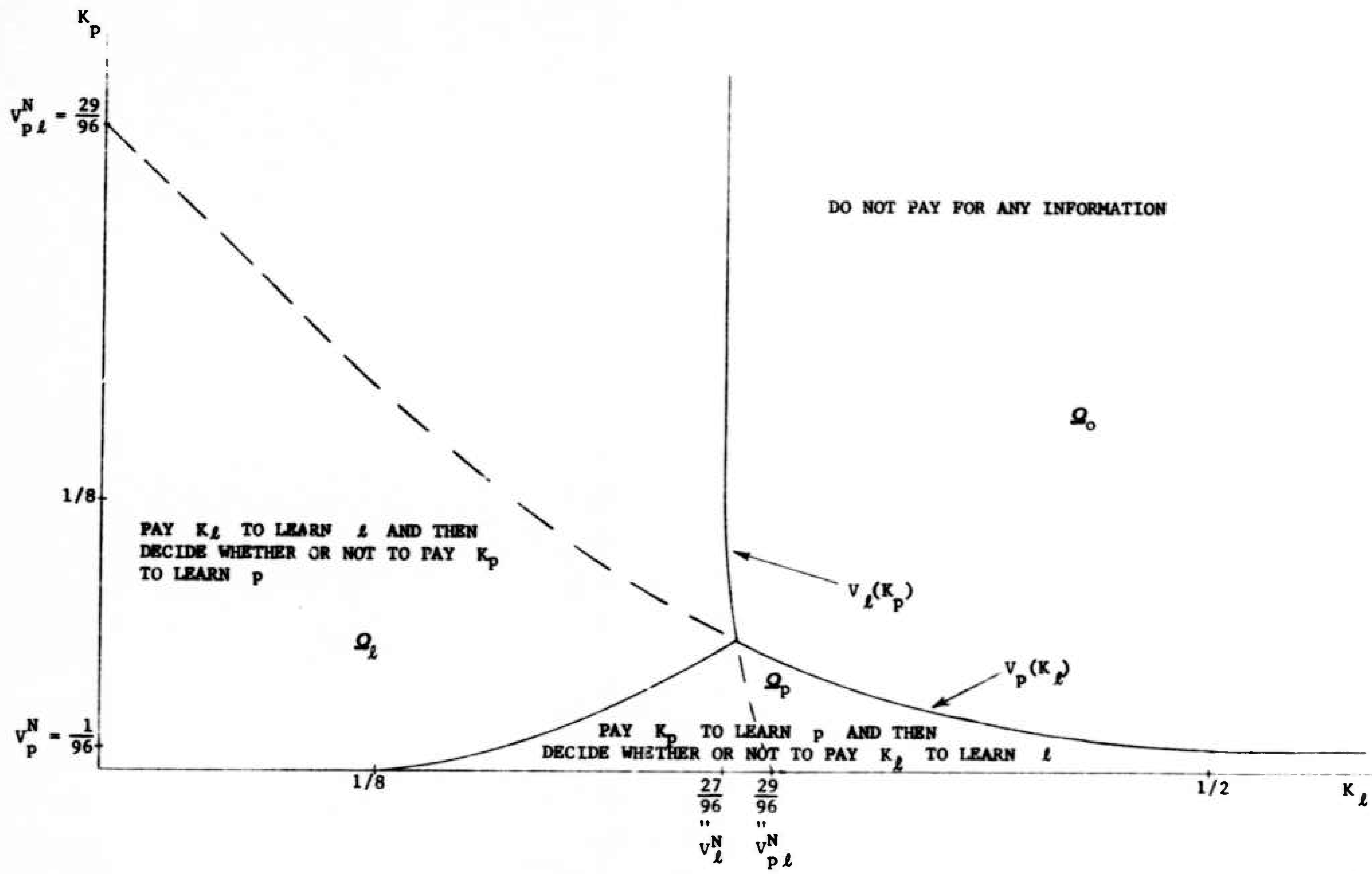


Figure 3. Decision regions for sequential information

these functions form boundaries of the decision regions. In this problem the option of buying information simultaneously is completely dominated by the other alternatives, even though there are pairs of prices such that buying both pieces of information simultaneously is preferable to buying either piece of information individually. A comparison of Figs. 2 and 3 shows that when information is available sequentially, we are willing to buy information at pairs of prices that were not advantageous when only individual and simultaneous information was available. Furthermore, there are pairs of prices such that our best decision when offered individual and simultaneous information is to pay $K_\ell$ to learn $\ell$, while out best initial decision when offered all types of information is to pay $K_p$ to learn $p$.

## General Properties of Sequential Information Problems with Additive, Certain Prices

When all observable prices are additive and certain, we can formulate a general sequential-information problem in terms of a set of state variables $(x_1,\ldots,x_m)$ and a set of observables $(y_1,\ldots,y_n)$ with the corresponding set of observable prices $(K_{y_1},\ldots, K_{y_n})$. When an observable is equal to one of the state variables it represents perfect information. However, by treating observables and state variables separately, we can also deal with imperfect information. Since the possibility of buying observables sequentially increases the value of both imperfect and perfect information, the value of imperfect information about a state variable can exceed the value of perfect, individual information about the same state variable.

As the preceding example demonstrates, the value of sequential information about an observable is, in general, a function of the prices of all of the other observables. Thus,

$$V_{y_i} = V_{y_i}(K_{y_1},\ldots,K_{y_{i-1}},K_{y_{i+1}},\ldots,K_{y_n})$$

The dependence of $V_{y_i}$ on the prices of the observables can be seen in the decision tree in Fig. 4. This tree shows all the information-purchasing decisions as well as the primary decision to choose a value for the control variable. The value of sequential information about $y_i$ is equal to the difference between the expected profit associated with branch $B_i$ (the branch where we first learn $y_i$ ) and the expected profit associated with branch $B_0$ (the branch where we do not learn any information. Using the tree in Fig. 4 we can find $V_{y_i}$ algebraically.

$$V_{y_i} = \underset{y_i}{E} \max \left\{ \begin{array}{l} \max\limits_{c} \underset{x_1}{E} \cdots \underset{x_m}{E} \pi(x_1,\ldots,x_m,c) \\ \max\limits_{j\neq i} \left( \underset{y_j}{E} \max \left\{ \begin{array}{l} \max\limits_{c} \underset{x_1}{E} \cdots \underset{x_m}{E} \pi(x_1,\ldots,x_m,c) \\ \max\limits_{k\neq i,j} (\underset{y_k}{E} \max \{\ldots\} - K_{y_k}) \end{array} \right\} - K_{y_j} \right) \end{array} \right\}$$

$$- \max_{c} \underset{x_1}{E} \cdots \underset{x_m}{E} \pi(x_1,\ldots,x_m,c)$$

This function has several properties which were demonstrated in the preceding example. Specifically we can show that as any price increases, $V_{y_i}$ must decrease or remain constant. However $V_{y_i}$ cannot decrease by more than the price increase. Thus,

$$\frac{\partial V_{y_i}}{\partial K_{y_j}} \in [-1,0]$$

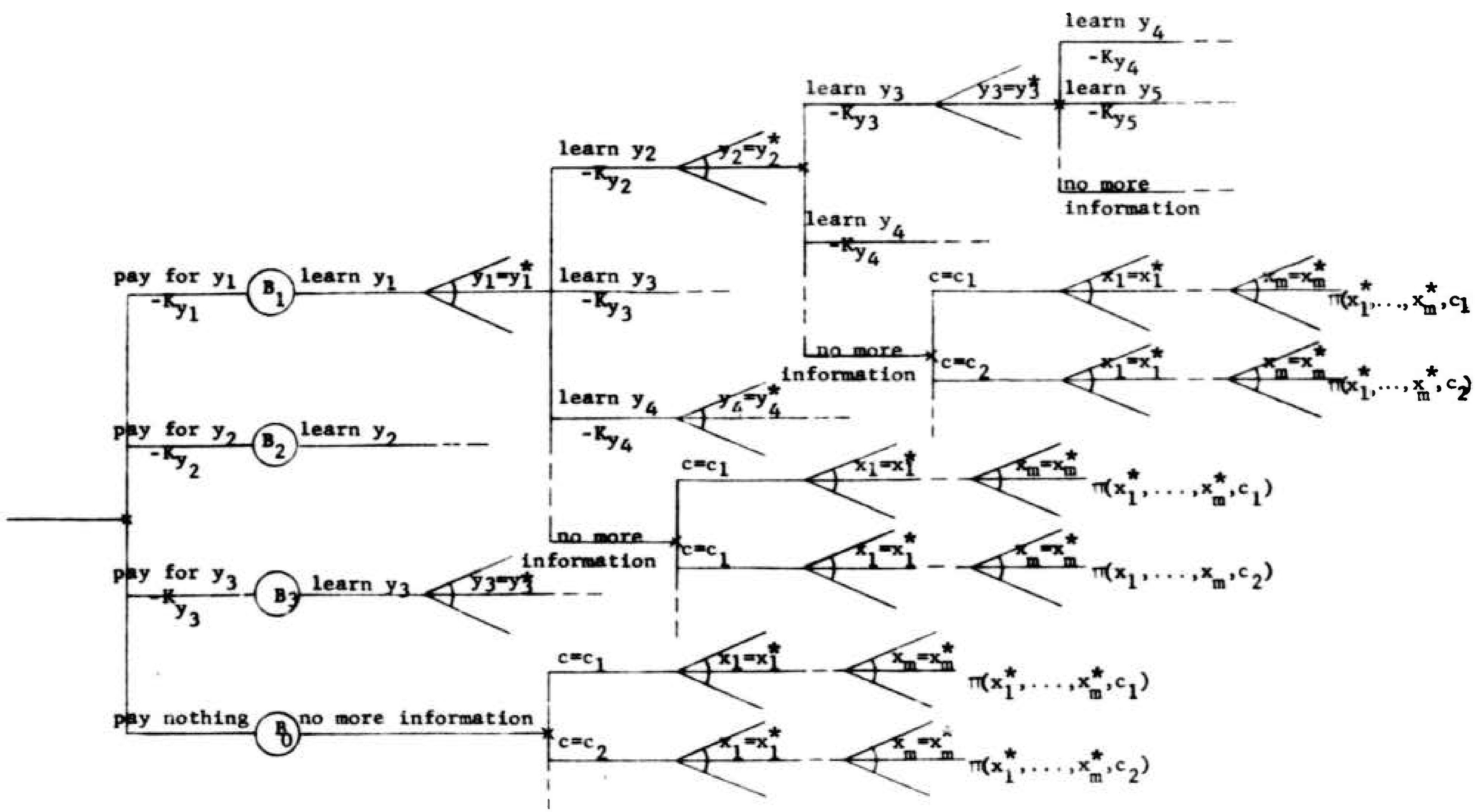


Figure 4. Sequential information decision tree for an expected profit decision maker

The proof of this statement [5] depends on the fact that all of the prices in the expression for $V_{y_i}$ are subtracted from some quantity. The derivative of $V_{y_i}$ becomes a series of expectations of quantities that are either minus one, zero, or some number between these limits.

It can also be shown that

$$\partial^2 V_{y_i} / \partial K_{y_j}^2 \geq 0$$

The proof [5] is based on the fact that, as $K_{y_j}$ increases, it becomes less desirable for the decision maker to pay for the $j^{th}$ observable regardless of the information he has already learned. In the expression for $V_{y_i}$ this means that there are fewer values of previously-learned observables for which the expectations include a term with $-K_{y_j}$. Thus $\partial V_{y_i} / \partial K_{y_j}$ becomes an expectation of more terms that are zero and fewer terms that are minus one, so the first derivative must increase or remain constant when $K_{y_j}$ increases.

The significance of these properties of the value of sequential information is that they allow $V_{y_i}$ to exceed the corresponding values of individual and simultaneous information for certain sets of observable prices. It is easy to see that $V_{y_i}$ must be at least as great as $V_{y_i}^N$ and $V_{y_i}^R$ for any set of observable prices.* When we buy information sequentially, we have the opportunity to buy just one observable or all of the observables. Thus we can always achieve the expected profits associated with individual and simultaneous information by making the

---

*Using $V_{y_i}^R(K_{y_1},\ldots,K_{y_{i-1}},K_{y_{i+1}},\ldots,K_{y_n})$, we can describe the price pairs where we will buy simultaneous information in terms of a function similar to $V_{y_i}$. We are willing to buy all the observables simultaneously when $K_{y_i} < V_{y_i}^R$.

proper set of sequential decisions.

The bidding example showed that, for certain sets of observable prices, $V_{y_i}$ can exceed both $V^N_{y_i}$ and $V^R_{y_i}$. This phenomenon, which makes it possible to buy sequential information at sets of prices that are not advantageous for individual or sequential purchases, occurs whenever learning $y_i$ can affect our decision to learn other observables. Suppose that for some set of observable prices our decision to buy additional information changes when we learn different values of $y_i$. Therefore, for some value of $y_i$, our best decision must be to buy additional information. For this value of $y_i$, our expected profit must be greater when we pay for some other observable, $y_j$, than when we refuse additional information. Therefore,

$$\underset{y_i}{D}\ \underset{y_j}{E}\ \max \begin{Bmatrix} \max\limits_{c} \underset{x_1}{E} \cdots \underset{x_m}{E}\ \pi \\ \max\limits_{k \neq i,j} \left( \underset{y_k}{E} \max \{\ldots\} - K_{y_k} \right) \end{Bmatrix} - K_{y_j}$$

$$> \underset{y_i}{D}\ \max_{c} \underset{x_1}{E} \cdots \underset{x_m}{E}\ \pi$$

Expanding this inequality and taking the expected value of both sides yields $V_{y_i} > V^N_{y_i}$. The proof that $V_{y_i}$ exceeds $V^R_{y_i}$ is more complicated but essentially similar [5].

## Observables with Uncertain Prices

If the prices of the observables are uncertain, we can assign prior probability density functions for each price. In general, the resulting distributions need not be independent, so our state of information is represented by a joint distribution on the state variables, observables,

and prices

$$\{x_1,\ldots,x_m,y_1,\ldots,y_n,K_{y_1},\ldots,K_{y_n}\,|\,\&\}$$

When we commit ourselves to paying for the $i^{th}$ observable, we learn both $y_i$ and $K_{y_i}$. After paying $K_{y_i}$ we must then decide whether to buy additional information.

The expected value of learning $y_i$ sequentially, when we are trying to maximize expected profit, and when the cost of information is uncertain, is

$$\underset{y_i}{E}\ \underset{K_{y_i}}{E}\ \max\left\{\begin{array}{l}\max\limits_{c}\ \underset{x_1}{E}\cdots\underset{x_m}{E}\ \pi(x_1,\ldots,x_m,c)\\[2ex] \max\limits_{j\neq i}\left(\underset{y_j}{E}\ \underset{K_{y_j}}{E}\left[\max\left\{\begin{array}{l}\max\limits_{c}\ \underset{x_1}{E}\cdots\underset{x_m}{E}\ \pi(x_1,\ldots,x_m,c)\\[1ex] \max\limits_{k\neq i,j}\left(\underset{y_j}{E}\ \underset{K_{y_j}}{E}\ [\ldots]\right)\end{array}\right\} - K_{y_j}\right]\right)\end{array}\right\}$$

$$-\max_{c}\ \underset{x}{E}\cdots\underset{x}{E}\ \pi(x_1,\ldots,x_m,c)$$

If $K_{y_j}$ is independent of $y_i$ and $K_{y_i}$, for all $i$ and $j$, this expression reduces to $V^N_{y_i}(\bar{K}_{y_1},\ldots,\bar{K}_{y_{i-1}},\bar{K}_{y_{i+1}},\ldots,\bar{K}_{y_n})$. In other words we can use the same decision rules for buying information that we would with certain prices, except we use the expected values of the prices instead of the prices themselves.

On the other hand, if $K_{y_j}$ is dependent on one of the other prices or the corresponding observables, we cannot characterize each uncertain price by its expected value. The maximizations required to determine the value of sequential information depend on quantities like $\underset{y_i}{D}\ \underset{K_{y_i}}{D}\ \underset{y_j}{E}\ K_{y_j}$, which in turn require a knowledge of the entire joint

density function for the state variables, observables, and prices. Since dependencies between these variables cannot generally be represented by a finite set of numbers, it is impossible to represent the decision rules in a finite-dimensional Euclidean space. However if we know the joint density function, we can determine the optimum strategy for buying information sequentially for that particular density function by looking at all of the ways we could buy the observables.

## Observables with Non-Additive Prices

If the observable prices are certain but not additive, it is necessary to define a large number of prices, one for each of the ways we could learn each observable. Some of the prices are:

$K_{y_i}$ = cost of learning $y_i$ when none of the other observables are known;

$K_{y_i|y_j}$ = cost of learning $y_i$ when $y_j$ is known;

$K_{y_i|y_jy_k\cdots y_r}$ = cost of learning $y_i$ when $y_j$ , $y_k$ , etc. are known;

$K_{y_iy_j\cdots y_r}$ = cost of learning $y_i$ , $y_j$ , etc. simultaneously when none of the observables are known and no additional information will be purchased.

Since each of these prices represents one dimension of a diagram such as that in Fig. 3, it is clear that we cannot hope to visualize the decision rules for buying information. However we can still describe the decision rules algebraically.

The decision rules can be simplified if we restrict the prices of the observables such that they can be described with relatively few numbers. One such restriction is to assume that the $i^{th}$ observable costs

$K_{y_i}$ if it is the first piece of information that we learn, and $\lambda_{y_i} K_{y_i}$ otherwise. The reduction factor $\lambda_{y_i}$ represents the savings that result from having previously set up an information-gathering process to learn a different observable. In this case we can visualize the decision rules in terms of a diagram, such as that in Fig. 3, with the boundary representing $V_{y_i}$ stretched along the $K_{y_i}$ axes by a factor of $(1/\lambda_{y_i})$ .

## Conclusions

We cannot regard the value of learning one observable by itself as the maximum that we would be willing to pay for that piece of information. When it is possible to buy additional information, the value of the first observable may increase. How much it increases depends on the prices of the other observables, so it is necessary to know the prices of all the observables before we can decide whether to buy one of them. The decision to buy a piece of information must take into account all of the ways we can purchase the information individually, simultaneously, and sequentially.

## References

[1] Howard, R.A., "Decision Analysis: Applied Decision Theory," Fourth International Conference on Operational Research, Boston, 1966.

[2] Howard, R.A., "Information Value Theory," IEEE Transactions on Systems Science and Cybernetics, Vol. SSC-2, No. 1, August 1906.

[3] Howard, R.A., "Value of Information Lotteries," IEEE Transactions on Systems Science and Cybernetics, Vol. SSC-3, June 1967.

[4] IEEE Transactions of Systems Science and Cybernetics, Special Issue on Decision Analysis, Vol. SCC-4, No. 3, September 1968.

[5] Miller, A.C., "The Value of Sequential Information," Ph.D. dissertation, Stanford University, 1973.

[6] Raiffa, H., and R.O. Schlaifer, "Applied Statistical Decision Theory," Harvard Graduate School of Business Administration, Harvard University, 1961.

[7] Raiffa, H., "Decision Analysis: Introductory Lectures on Choices under Uncertainty," Addison-Wesley, 1968.